**MATH 217 W24 - LINEAR ALGEBRA, Section 001 (Dr. Paul Kessenich)**
**Homework Set Part B due Thurs, Feb 8 at 11:59pm**
**Zhengyu James Pan (jzpan@umich.edu)**

1. Let $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^n$ be a linear transformation. As on HW 3, we define $T^k$ to be the $k$-fold composition of $T$ with itself. Let $A$ be the standard matrix of $T$, by which we mean the unique $n \times n$ matrix such that $T(\vec{x}) = A\vec{x}$ for all $\vec{x} \in \mathbb{R}^n$.

(a) Prove that for all $k$, the standard matrix for $T^k$ is the matrix $A^k$. [Hint: induction works nicely.]

**Solution:** We are given that the standard matrix $A^{(1)}$ represents the transformation $T^{(1)}$. Assume that the transformation $T^n$ can be represented by the standard matrix $A^n$. We know by a theorem on the worksheets that the standard matrix of two linear transformations, both from $\mathbb{R}^n \to \mathbb{R}^n$, is equal to the product of their respective standard matrices. Then $(T^n \circ T)(x) = A^n A\vec{x}$. This is equal to $T^{n+1}(x) = A^{n+1}\vec{x}$. So by induction, $T^k(\vec{x}) = A^k\vec{x}$ for all $\vec{x} \in \mathbb{R}^n$ and $k \in \mathbb{N}$.

(b) We define $T$ to be nilpotent if there exists some $k \in \mathbb{N}$ such that $T^k$ is the zero transformation. Prove that if $T$ is nilpotent, then $A$ is not invertible.

**Solution:** We will prove the contrapositive. Assume $A$ is invertible. We know zero transformation is surjective, so it is noninvertible, and We know by part (a) that the standard matrix of $T^k$ is $A^k$ for natural number $k$. By problem 5c on Worksheet 7, the inverse of $A^k$ is $(A^{-1})^k$. So $A^k$ can never be noninvertible, and is thus never the zero transformation. Thus if $A$ is invertible, then $T$ is not nilpotent, so the original statement is true.

(c) Prove that if $T$ is nilpotent, then $A - I_n$ is invertible. [Hint: try multiplying out $(A - I_n)(-I_n - A - A^2 - \cdots - A^{k-1})$ and see what you get.]

**Solution:** Let $k \in \mathbb{N}$ such that $T^k$ is the zero transformation.
Expanding $(A - I_n)(-I_n - A - A^2 - \cdots - A^{k-1})$:

$$
\begin{aligned}
&= -A\left(I_n + A + \cdots + A^{k-1}\right) + I_n\left(I_n + A + \cdots + A^{k-1}\right) &&\text{(distributivity)}\\
&= -\left(A + A^2 + \cdots + A^{k-1} + A^k\right) + \left(I_n + A + \cdots + A^{k-1}\right) &&\text{(distributivity)}\\
&= \left(I_n + A + \cdots + A^{k-1}\right) - \left(A + A^2 + \cdots + A^{k-1} + 0_{n\times n}\right) &&(A^k = 0_{n\times n})\\
&= I_n
\end{aligned}
$$

Swapping the order of multiplication,

$$
\begin{aligned}
(-I_n &- A - A^2 - \cdots - A^{k-1})(A - I_n)\\
&= -\left(I_n + A + \cdots + A^{k-1}\right)A + \left(I_n + A + \cdots + A^{k-1}\right)I_n &&\text{(distributivity)}\\
&= -\left(A + A^2 + \cdots + A^{k-1} + A^k\right) + \left(I_n + A + \cdots + A^{k-1}\right) &&\text{(distributivity)}\\
&= \left(I_n + A + \cdots + A^{k-1}\right) - \left(A + A^2 + \cdots + A^{k-1} + 0_{n\times n}\right) &&(A^k = 0_{n\times n})\\
&= I_n
\end{aligned}
$$

So the inverse of $A - I_n$ is $-(I_n + A + A^2 + \cdots + A^{k-1})$, and $A - I_n$ is invertible by definition.

2. Let $V$ be any vector space, and let $S$ be any set. Let $\mathcal{F}(S, V)$ denote the set of all functions from $S$ to $V$ . (Note: we are not assuming $S \subseteq V$ here, just that $S$ is some set. $S$ is not assumed to be a vector space, but it could be. Similarly, the functions in $F(S, V)$ are not assumed to be linear transformations, although it is possible that some of them might be.)

For any functions $f, g \in \mathcal{F}(S, V)$ we can define their sum to be the function $f + g$ given by the formula $(f + g)(s) = f(s) + g(s)$, where $s$ is any element in $S$. Similarly, for any scalar $c \in R$ and any function $f \in \mathcal{F}(S, V)$ we define the function $cf$ to be given by the formula $(cf)(s) = c(f(s))$ for all $s \in S$.

(a) Prove that $\mathcal{F}(S, V)$ is a vector space. Note: For this problem you must explicitly prove that each of the vector space properties VS1-8 from Worksheet 6 is true. (These proofs should be very short but are not skippable.)

**Solution:** Let arbitrary $a, b \in \mathbb{R}$, arbitrary $f, g, h \in \mathcal{F}(S, V)$. Note that $+_{\mathcal{F}(S,V)}$ borrows the qualities of $+_V$ (the summation operation of $V$) through the definition; namely associativity and commutativity.

VS-1: True, $(f + g) + h = (f(s) + g(s)) + h(s) = f(s) + (g(s) + h(s)) = f + (g + h)$ by additive associativity of vector space $V$.
VS-2: $f + g = f(s) + g(s) = g(s) + f(s) = g + f$ by addititve commutativity of vector space $V$.
VS-3: True, $f(s) = 0_V$ satisfies this property. $g + f = g(s) + 0_V = g$
VS-4: True, for all values of $f(s)$, such a $-f(s)$ exists since $V$ is a vector space. So the function $-f$ exists as well.
VS-5: True, $a(f + g) = a(f(s) + g(s)) = af(s) + ag(s) = af + ag$ by distributivity of the vector space $V$.
VS-6: True, $(a + b)f = (a + b)f(s) = af(s) + bf(s) = af + bf$. by scalar multiplicative distributivity of vector space $V$.
VS-7: True, $a(bf) = a(bf(s)) = (ab)f(s) = (ab)f$ by scalar multiplicative associativity of vector space $V$.
VS-8: True, $1f = 1f(s) = f(s) = f$ by the unitary law of vector space V.

(b) Is $0_{\mathcal{F}(S,V)}$ the same element as $0_V$ ? If not, explain how they are different.

**Solution:** No. $0_{\mathcal{F}(S,V)}$ maps any element of the set $S$ to $0_V$, while $0_V$ is only a vector in the space $V$. $0_{\mathcal{F}(S,V)}$ is a function and can take an input, while $0_V$ cannot take an input like a function.

(c) We could similarly define $\mathcal{F}(V, S)$ to be the set of all functions from $V$ to $S$. Would $\mathcal{F}(V, S)$ also a vector space? Why or why not?

**Solution:** Not necessarily. We used the vector space properties of image $V$ to prove the vector space axioms for $\mathcal{F}(S, V)$. However, when arbitrary set $S$ is the image, those properties do not necessarily apply.

(d) The familiar vector spaces $\mathcal{P}$, $\mathcal{P}_n$ and $\mathcal{C}^\infty$ (all from Worksheet 6) are all subsets of $\mathcal{F}(S, V)$ for some $S$ and $V$ . What are $S$ and $V$ for each of these functions?

**Solution:** All of these vector spaces are composed of functions which map from $\mathbb{R} \to \mathbb{R}$.

3. Let $\mathcal{P}$ be the vector space of all polynomial functions from $\mathbb{R}$ to $\mathbb{R}$ in the variable $t$, and for each $n \in \mathbb{N}$, let $\mathcal{P}_n$ be (as usual) the subset of $\mathcal{P}$ consisting of all polynomial functions of degree at most $n$. (We already know that $\mathcal{P}_n$ is also a vector space.) Also let $T : \mathcal{P} \to \mathcal{P}$ be the map defined by $T(p)(t) = p'(t) + p(0)$ for each $p \in \mathcal{P}$ and for all $t \in \mathbb{R}$.

(a) Show that $T$ is a linear transformation.

**Solution:** Note that the derivative is linear by problem 5 of worksheet 6. Although the current domain and codomain are $\mathcal{P}$ and not $\mathcal{C}^\infty$, the derivative is closed in $\mathcal{P} \subset \mathcal{C}^\infty$, so it is still linear. First, we show $T$ respects addition:

$$
\begin{aligned}
T(p+q)(t) &= (p+q)'(t) + (p+q)(0) && \text{(definition of } T) \\
&= p'(t) + q'(t) + p(0) + q(0) && \text{(linearity of derivative)} \\
&= p'(t) + p(0) + q'(t) + q(0) && \text{(associativity and commutativity of } +) \\
&= T(p) + T(q) && \text{(definition of } T(p))
\end{aligned}
$$

Next, we show $T$ respects scalar multiplication. Let $c \in \mathbb{R}$.

$$
\begin{aligned}
T(cp) &= (cp)'(t) + (cp)(0) && \text{(definition of } T) \\
&= c(p'(t)) + cp(0) && \text{(linearity of derivative)} \\
&= c(p'(t) + p(0)) && \text{(distributivity of scalar multiplication in vector space } \mathcal{P}) \\
&= cT(p)
\end{aligned}
$$

Thus $T$ is linear.

(b) Let $n \in \mathbb{N}$, and let $T_n : \mathcal{P}_n \to \mathcal{P}_n$ be defined by $T_n(p)(t) = p'(t) + p(0)$, so that $T_n$ is just $T$ with both domain and codomain restricted to $\mathcal{P}_n$. Is $T_n$ injective? Is $T_n$ surjective?

**Solution:** $T$ is not surjective in $\mathcal{P}_n$: no $p \in \mathcal{P}_n$ satisfies $T(p) = x^n$. $T$ is not injective in $\mathcal{P}_n$: $T(2x) = 2 = T(2)$.

(c) Is $T$ injective? Is $T$ surjective?

**Solution:** $T$ is surjective: for any polynomial $p(x) \in \mathcal{P}$, let $q = \int p(x)\,dx$ with any constant term. We know this is in $\mathcal{P}$ by the power rule for antiderivatives. Then $T(q) = p(x)$, so $T$ is surjective. However, $T$ is not injective: $T(2x) = 2 = T(2)$.

4. We denote by $\mathbb{R}^{n\times n}$ the vector space of all $n \times n$ matrices. Let $A$ be an $n \times n$ matrix, and define the function $L_A : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ by $L_A(B) = AB$ for all $B \in \mathbb{R}^{n\times n}$. (Note carefully: this is *not* the same function as $T_A$. While both can be described informally as "multiplication by $A$", the two functions $L_A$ and $T_A$ have different domains and codomains. Make sure you understand this distinction before beginning to work on this problem!)

(a) Show that $L_A$ is a linear transformation.

**Solution:** Let $B, C \in \mathbb{R}^{n\times n}$, and $k \in \mathbb{R}$. We first show $L_A$ respects addition:

$$
\begin{aligned}
L_A(B+C) &= A(B+C)           &&\text{(definition of } L_A) \\
&= AB + AC                    &&\text{(distributivity of matrix multiplication)} \\
&= L_A(B) + L_A(C)
\end{aligned}
$$

Next, we show $L_A$ respects scalar multiplication.

$$
\begin{aligned}
L_A(kB) &= A(kB)              &&\text{(definition of } L_A) \\
&= (Ak)B                      &&\text{(scalar associativity of matrix multiplication)} \\
&= (kA)B                      &&\text{(scalar commutativity of matrices)} \\
&= k(AB) \\
&= kL_A(B)                    &&\text{(definition of } L_A)
\end{aligned}
$$

So $L_A$ is linear.

(b) Show that the matrix $A$ is invertible if and only if the linear transformation $L_A$ is invertible.

**Solution:** Assume the linear transformation $L_A$ is invertible. Then there exists $L_A^{-1}$ such that $L_A \circ L_A^{-1} = L_A^{-1} \circ L_A$ is the identity function. Let $C \in \mathbb{R}^{n\times n}$ such that $C = L_A^{-1}(I_n)$. Then $A(C) = L_A\left(L_A^{-1}(I_n)\right) = I_n$. So $A$ is invertible by corollary 2.19 from the Theory of Linear Algebra handout.

Assume $A$ is invertible. Let its inverse be $C$. Then $L_A(L_C(B)) = A(CB) = (AC)B = I_n B$ is the identity function. Similarly, $L_C(L_A(B)) = C(AB) = (CA)B = I_n B$ is the identity function. So $L_A$ is invertible with inverse $L_C$.

Now let $\mathcal{F}$ be the set of all functions from $\mathbb{R}^{n\times n}$ to $\mathbb{R}^{n\times n}$, and define the function $L : \mathbb{R}^{n\times n} \to \mathcal{F}$ by $L(A) = L_A$.

(c) Show that $L$ is injective.

**Solution:** Let $A_1, A_2 \in \mathbb{R}^{n\times n}$ be such that $L(A_1) = L(A_2)$. Then $L_{A_1}(I_n) = A_1 I_n = A_1$ is equal to $L_{A_2} = A_2 I_n = A_2$. So $A_1 = A_2$, and $L$ is injective.

(d) Is $L$ surjective? Be sure to justify your claim.

**Solution:** No, $L$ cannot map to functions which are nonlinear by part (a), even though such functions assuredly exist for $\mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$.

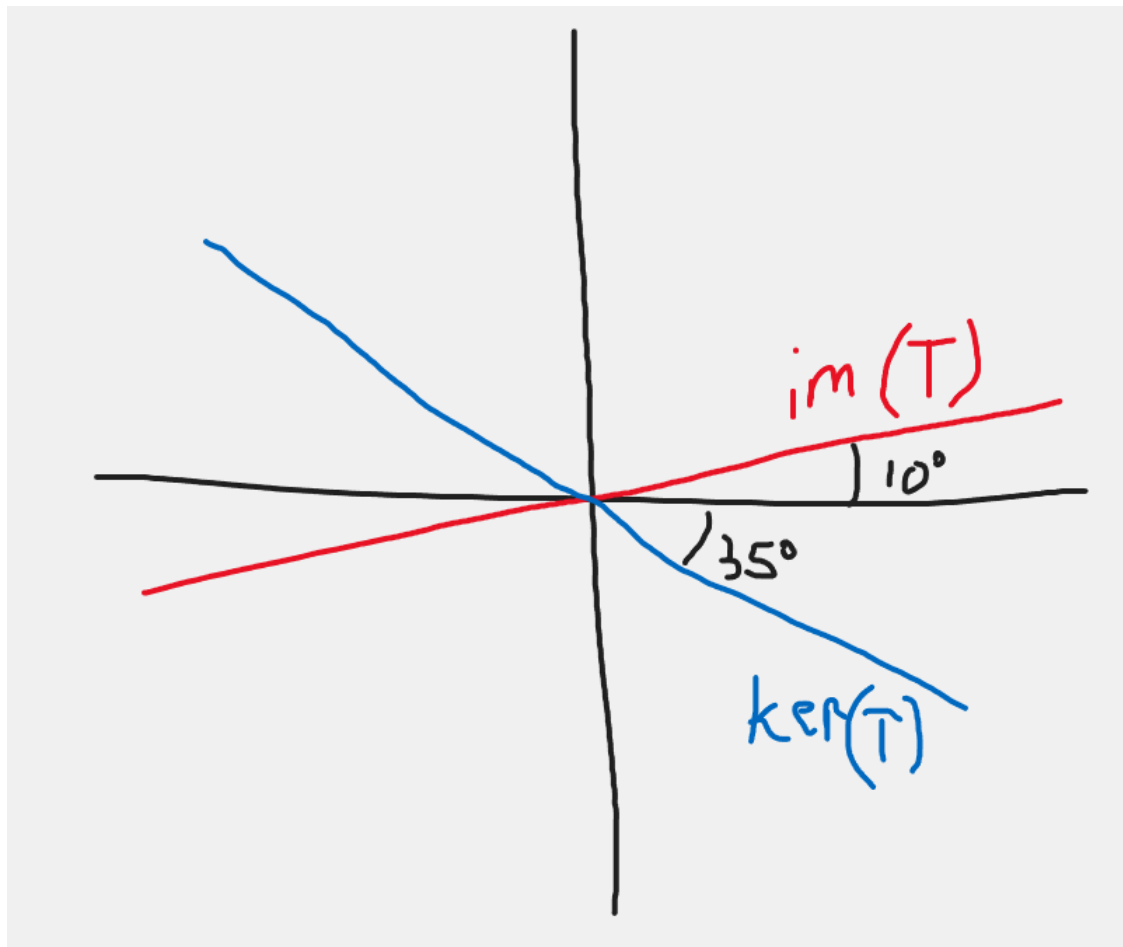5. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation defined as follows:

$$T = \mathrm{Rot}_{-80°} \circ \mathrm{Proj}_y \circ \mathrm{Rot}_{35°},$$

where $\mathrm{Rot}_\theta$ is counter-clockwise rotation by $\theta$, and $\mathrm{Proj}_y$ is projection onto the $y$-axis.

(a) Sketch im($T$) in $\mathbb{R}^2$. Indicate the angle between im($T$) and the $x$-axis.

(b) Sketch ker($T$) in $\mathbb{R}^2$. Indicate the angle between ker($T$) and the $x$-axis.

**Solution:**



The image is the image of the y-axis rotated by 80 degrees clockwise, which forms an angle of 10 degrees with the x-axis.

The kernel of the rotation is the origin. The kernel of the projection is the x-axis. So the line 35 degrees clockwise of the x-axis is the kernel of the compositions.

(c) Let $T_{\phi,\theta} := \mathrm{Rot}_\theta \circ \mathrm{Proj}_y \circ \mathrm{Rot}_\phi$. For which $\phi$ and $\theta$ is im($T_{\phi,\theta}$) = ker($T_{\phi,\theta}$)?

**Solution:** Using the logic from the previous part, the y-axis rotated by $\theta$ counter-clockwise should match the x-axis rotated by $\phi$ clockwise. So $\theta + 90° = -\phi$.